I. AMENDMENTS

Specification

Please amend the paragraph bridging pages 30-31 as follows:

A number of genome databases are available for comparison, including, for example, a substantial portion of the human genome is available as part of the Human Genome Sequencing Project (J. Roach, <u>at uniform resource language ("URL")</u> <u>"http:(//)weber.u.Washington.edu/~roach/human_genome_progress 2.html"</u> ~~http://weber.u.Washington.edu/~roach/human_genome_progress 2.html~~). In addition, at least twenty-one genomes have been sequenced in their entirety, including, for example, *M. genitalium*, *M. jannaschii*, *H. influenzae*, *E. coli*, yeast (*S. cerevisiae*), and *D. melanogaster*. Significant progress has also been made in sequencing the genomes of model organism such as mouse, *C. elegans*, and *Arabadopsis sp*. Several databases containing genomic information annotated with some functional information are maintained by different organizations, and are accessible via the internet, for example, <u>at URL "http:(//)wwwtigr.org/tdb"; at URL</u> <u>"http:(//)www(dot)genetics.wisc.edu"; at URL "http:(//)genome-www.stanford.edu/~ball"; at ULR</u> <u>"http:(//)hiv-web.lanl.gov"; at URL "http:(//)www(dot)ncbi.nlm.nih.gov"; at URL</u> <u>"http:(//)www(dot)ebi.ac.uk"; at URL "http:(//)Pasteur.fr/other/biology; and at URL</u> <u>"http:(//)www(dot)genome.wi.mit.edu"</u> ~~http://wwwtigr.org/tdb; http://www.genetics.wisc.edu;~~ ~~http://genome-www.stanford.edu/~ball; http://hiv-web.lanl.gov; http://www.ncbi.nlm.nih.gov;~~ ~~http://www.ebi.ac.uk; http://Pasteur.fr/other/biology; and http:// www.genome.wi.mit.edu~~.

Please amend the paragraph pages 31-32 (Note - underlining of references in original) as follows:

One example of a useful algorithm is BLAST and BLAST 2.0 algorithms, which are described by Altschul et al. (Nucleic Acids Res. 25:3389-3402, 1977; J. Mol. Biol. 215:403-410, 1990, each of which is incorporated herein by reference). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (<u>at URL</u> <u>"www(dot)ncbi.nlm.nih.gov"</u> ~~http://www.ncbi.nlm.nih.gov~~). This algorithm involves first

identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul et al., *supra*, 1977, 1990). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always >0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=4 and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectations (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff, Proc. Natl. Acad. Sci., USA 89:10915, 1989) alignments (B) of 50, expectation (E) of 10, M=5, N=-4, and a comparison of both strands.

Please amend the paragraph bridging pages 32-32 (Note - underlining of reference in original) as follows:

The BLAST programs identify homologous sequences by identifying similar segments, which are referred to herein as "high-scoring segment pairs," between a query amino or nucleic acid sequence and a test sequence which is preferably obtained from a protein or nucleic acid sequence database. High-scoring segment pairs are preferably identified (aligned) by means of a scoring matrix, many of which are known in the art. Preferably, the scoring matrix used is the BLOSUM62 matrix (Gonnet et al., Science 256:1443-1445, 1992; Henikoff and Henikoff, Proteins 17:49-61, 1993, each of which is incorporated herein by reference). Less preferably, the

In re Application of:
    Lee and McPherron
Application No.: 09/640,887
Filed: August 17, 2000
Page 4

PATENT
Attorney Docket No.: JHU1120-13

PAM or PAM250 matrices may also be used (Schwartz and Dayhoff, eds., "Matrices for Detecting Distance Relationships: Atlas of Protein Sequence and Structure" (Washington, National Biomedical Research Foundation 1978)). BLAST programs are accessible through the U.S. National Library of Medicine, for example, at URL "www(dot)ncbi.nlm.nih.gov" at ~~www.ncbi.nlm.nih.gov~~.